EXHIBIT 6

CHAPTER

4

Demographic Adjustments to WAIS–IV/WMS–IV Norms

James A. Holdnack and Larry G. Weiss
Pearson Assessment, San Antonio, Texas, USA

INTRODUCTION

Clinicians frequently evaluate patients from diverse socioeconomic and cultural backgrounds. When evaluating patients from minority or economically disadvantaged environments, the clinician needs to consider the extent to which cultural, educational, financial, and other environmental factors, such as access to healthcare, impact performance on cognitive tests. An accurate diagnosis requires the clinician to discern the degree to which an achieved low score represents the effects of a disease process rather than the impact of cultural, educational, and economic factors. Similarly, the extent to which a high level of educational achievement and financial status impact test performance should also be considered, particularly in the presence of possible cognitive loss or decline where average scores may indicate a loss of function. In these evaluations, the clinician attempts to determine if an achieved score represents an *expected* level of performance for that *individual* based on his or her background characteristics.

Standardized scores describe the rank order of an examinee's performance in comparison to individuals of similar age; however, the standardized score does not indicate whether the level of performance is atypical for individuals with similar background characteristics. A very low score reflects low cognitive functioning currently; however, it may not represent an unexpectedly low score in an examinee with a long history of low cognitive ability. Therefore, scores need to be

WAIS-IV, WMS-IV, and ACS.
DOI: http://dx.doi.org/10.1016/B978-0-12-386934-0.00004-3

evaluated within the context of the individual's background. If an examinee obtains a score of 90, it may represent a good, average, or poor performance for that person. For example, in a healthy examinee with 8 years of education and a history of low but stable occupational and educational performance, a score of 90 may be considered within expected limits. However, in an examinee with 21 years of education employed in a highly intellectual field, a score of 90 might represent an unexpectedly low performance. The difficulty of interpreting performance is compounded by the fact that cognitive skills vary considerably in the degree to which background factors impact test performance. Therefore, *expected* level of performance cannot accurately be determined without statistical data. At the extremes of a distribution (e.g., very high and very low education levels), the clinician can roughly estimate whether a score appears to be unexpectedly high or low; however, as background parameters approach the median it becomes increasingly difficult to accurately determine the level of expected performance. The use of demographic variables to estimate abilities improves the accuracy of estimation. It is important to note that interpreting an examinee's performance relative to their background characteristics is done *only* to identify an unexpected level of performance; that is to determine if a change has occurred from an estimated prior level of ability. Comparing a patient's performance to individuals with similar background characteristics should not be used to determine his or her ability to adequately function in general societal contexts or in place of standard scores in evaluating current performance (e.g., special education evaluations, disability due to chronic medical or psychiatric condition).

Ideally, current performance data is compared to previous results (i.e., serial assessment). In the absence of previous test data, the clinician can evaluate the current level of performance in relation to the expected performance given the patient's psychosocial history. An unexpectedly low score *potentially* represents a decline in cognitive functioning. In this situation, the clinical hypothesis involves a general or specific *loss* in cognitive functioning such as: *Does the patient have a significant loss in memory*; *Is there a loss in cognitive functioning following a moderate traumatic brain injury*; or *Is the patient's verbal ability impaired following cerebrovascular accident?* While it may be possible to determine if a patient's performance is higher than expected, clinical hypotheses rarely evaluate improved cognition without conducting serial assessments (see Chapter 6). Understanding the extent to which background factors relate to cognitive functioning is necessary for accurate application and interpretation of demographically adjusted data.

BACKGROUND FACTORS ASSOCIATED WITH TEST PERFORMANCE

Multiple independent and inter-related factors influence obtained test scores. An innumerable number of potential factors may influence test performance. The ability to detect changes in brain function in patients via changes in cognition is key in determining which variables to include in demographic adjustments. For example, if gender results in significantly higher scores for one group or another, this difference in performance could mask a change in functioning or result in misidentification of impairment when none is present. Specifically, if men are superior at visual–perceptual functioning compared to women, then the normative data based on combined male–female performance will yield higher scores for men. Subsequently, if men have an average or low average score using population norms, this *could* represent lower than expected performance. The degree that such differences may mask unexpectedly low performance depends on the effect size of the difference. The larger the effect size, the greater the clinical implications for accurately identifying cognitive deficits. The ACS demographic adjusted norms incorporate three demographic variables that have been used in research and clinical practice for decades: sex, education, and race/ethnicity (Heaton, Grant, & Matthews 1991). The relationship of each variable with WAIS–IV and WMS–IV performance will be briefly reviewed.

Sex Effects

Research comparing cognitive performance between men and women has identified differences in multiple domains. There is no consistent sex advantage in general cognition; in some areas, men outperform women, and in others, women outperform men. There is a large body of research on gender differences not only in cognition but also in brain morphology and function. While sex differences are observed on many tasks, not all differences are relevant to the current discussion. This chapter focuses on sex differences that could impact interpretation of results in neuropsychological assessments.

In general, men outperform women on tests of general cognitive functioning, perceptual reasoning, working memory (Heaton, Taylor, & Manly, 2003; van der Sluis et al., 2006) and verbal reasoning (Heaton et al., 2003). Women perform better than men on processing speed (Camarata & Woodcock, 2006; Heaton et al., 2003; Longman, Saklofske, & Fung, 2007; van der Sluis et al., 2006). These results are important for the WAIS–IV, as these skills are core to the WAIS–IV measurement

model. Within these broad domains, specific abilities can yield similar or quite different results. For example, men outperform women on perceptual reasoning tests, mental rotation tasks (Herlitz, Airaksinen, & Nordstrom, 1999), and spatial processing (Gur et al., 2012); however, women are better at face detection (McBain, Norton, & Chen, 2010). Men score higher on verbal reasoning measures, but women are better at verbal productivity (Herlitz et al., 1999). On working memory tasks, some research shows an advantage for men on verbal working memory (Heaton et al., 2003) while other studies show better visual spatial working memory but not verbal working memory (Lejbak, Crossley, & Vrbancic, 2011). Women are faster on measures of cognitive processing speed but men are faster on simple motor and sensorimotor speed tasks (Gur et al., 2012). General statements about superior functioning for one sex over the other do not apply to all skills within a domain. Therefore, the degree to which tests within domain measure different aspects of the construct in question impacts the degree and type of sex effects observed.

Differences in cognitive skills relevant to the WAIS−IV do not necessarily apply to the WMS−IV. The WMS−IV assesses auditory and visual, episodic, declarative memory functions. Research generally shows that women outperform men on many but not all types of memory functions. Women outperform men (Ragland, Coleman, Gur, Glahn, & Gur, 2000; Trahan & Quintana, 1990; Weirich, Hoffmann, Meissner, Heinz, & Bengner, 2011) and girls outperform boys (Gur et al., 2012; Kramer, Delis, Kaplan, O'Donnell, & Prifitera, 1997) on verbal learning tasks. Women show a particular advantage in the acquisition of new information rather than in maintaining acquired information (Krueger & Salthouse, 2010). The female verbal memory advantage may be related to differences in the pattern of blood flow in the temporal lobes of women versus men (Ragland et al., 2000). While single word verbal learning tasks show consistent female advantage, another verbal learning paradigm, verbal paired associates, does not demonstrate the same advantage for women (Trahan & Quintana, 1990). On visual memory tasks, inconsistent gender effects are identified with no gender effects observed on memory for visuospatial information and abstract objects (Herlitz et al., 1999; Trahan & Quintana, 1990), and females outperforming men on memory for concrete, nameable objects (Herlitz et al., 1999) and memory for faces (Lewin, Wolgers, & Herlitz, 2001). On the WMS−III, women performed significantly better than men on all the auditory and visual memory indexes (Heaton et al., 2003).

Table 4.1 presents descriptive statistics for women and men on the core WAIS−IV indexes and subtests. Performance was significantly different between men and women on all WAIS−IV index scores with a male advantage for VCI, PRI, and WMI; and a female advantage on

TABLE 4.1 WAIS–IV Core Index and Subtest Mean Scores, by Sex

	Male		**Female**		**Between Group Differences**	
Score	**Mean**	**SD**	**Mean**	**SD**	**Effect Size**	***p***
Verbal Comprehension Index	101.8	15.4	98.4	14.5	0.23	< 0.001
Perceptual Reasoning Index	101.8	15.6	98.3	14.2	0.24	< 0.001
Working Memory Index	101.8	15.5	98.4	14.3	0.23	< 0.001
Processing Speed Index	97.7	14.5	102.1	15.1	−0.30	< 0.001
Full Scale IQ	101.2	15.3	98.9	14.6	0.15	< 0.001
Vocabulary	10.1	3.1	10.0	2.9	0.05	> 0.05
Similarities	10.2	2.9	9.9	2.9	0.11	< 0.05
Information	10.7	3.2	9.4	2.8	0.45	< 0.001
Block Design	10.5	3.1	9.6	2.9	0.28	< 0.001
Matrix Reasoning	10.2	3.2	10.0	3.0	0.07	> 0.05
Visual Puzzles	10.4	3.2	9.7	2.8	0.25	< 0.001
Digit Span	10.2	3.1	9.9	3.0	0.08	> 0.05
Arithmetic	10.5	3.1	9.6	2.8	0.32	< 0.001
Coding	9.4	2.9	10.5	3.0	−0.39	< 0.001
Symbol Search	9.8	3.1	10.3	3.1	−0.15	< 0.001

PSI. The effect sizes are small ranging from 0.15 (FSIQ) to −0.30 (PSI). Greater variability is observed in sex differences at the subtest level. On VCI measures, scores on Similarities and Information were significantly higher for males than female but no difference was observed on Vocabulary. PRI subtests show a significant advantage for males on Block Design and Visual Puzzles but not on Matrix Reasoning. Auditory working memory (WMI) scores show a male advantage on Arithmetic but not on Digit Span. Women obtain significantly higher scores on both PSI subtests with the greatest difference observed on Coding compared to Symbol Search. While statistically significant differences are observed between men and women on WAIS–IV index and subtest scores, the small effect sizes indicate limited clinical significance of the differences. At the index level, the largest effect of −0.3, a difference of 5 standard score points, is similar to the standard error of measurement for most indexes. At the subtest level, the largest effect

is −0.39 or about 1 scaled score difference. This is also within the standard error of measurement for these measures.

Table 4.2 provides descriptive statistics for male versus female performance on the WMS–IV. Women show significantly higher scores on Immediate, Delayed, and Auditory Memory indexes and men obtain higher Visual Working Memory index scores. There was no gender difference observed on the Visual Memory Index. These results differ from those observed on the WMS–III for Visual Memory (Wechsler, 2007), likely due to the composition of the index. On the WMS–III, the visual memory measures included face recognition and verbal–visual

TABLE 4.2 WMS–IV Core Index and Subtest Mean Scores, by Sex

	Male		Female		Between Group Differences	
Score	**Mean**	**SD**	**Mean**	**SD**	**Effect Size**	***p***
Immediate Memory Index	99.1	14.7	101.0	15.1	−0.13	<0.05
Delayed Memory Index	99.0	15.1	101.0	14.8	−0.13	<0.05
Auditory Memory Index	98.2	14.8	101.7	15.0	−0.24	<0.001
Visual Memory Index	100.5	15.3	99.8	14.6	0.05	>0.05
Visual Working Memory Index	101.3	15.4	98.8	14.7	0.17	<0.05
Logical Memory I	9.8	3.1	10.2	2.9	−0.16	<0.01
Logical Memory II	9.6	3.1	10.3	2.9	−0.25	<0.001
Verbal Paired Associates I	9.7	2.9	10.3	3.1	−0.19	<0.001
Verbal Paired Associates II	9.7	3.0	10.3	3.1	−0.19	<0.01
Visual Reproduction I	10.2	3.1	10.0	3.0	0.06	>0.05
Visual Reproduction II	10.2	3.2	9.9	2.9	0.10	>0.05
Designs I	10.0	3.0	10.1	3.0	−0.03	>0.05
Designs I Content	9.9	3.1	10.2	3.0	−0.09	>0.05
Designs I Spatial	10.2	3.1	9.8	2.9	0.12	>0.05
Designs II	10.1	3.0	10.0	3.0	0.03	>0.05
Designs II Content	9.9	3.0	10.1	3.0	−0.08	>0.05
Designs II Spatial	10.4	3.0	9.6	2.9	0.27	<0.05
Spatial Addition	10.2	3.1	9.8	2.9	0.14	<0.05
Symbol Span	10.1	3.0	10.0	3.0	0.04	>0.05

association memory; on WMS−IV the visual memory measures require memory for visual detail and visual−spatial memory. The effect sizes for index level differences are quite small on WMS−IV, ranging from −0.24 (Auditory Memory Index) to 0.17 (Visual Working Memory Index). At the subtest level, women obtained higher scores on Logical Memory and Verbal Paired Associates immediate and delayed recall. Men obtained higher scores on delayed memory for spatial locations and spatial working memory. Men did not show an advantage on memory or working memory for visual details. The effect sizes on WMS−IV subtests range from −0.25 (Logical Memory II) to 0.27 (Spatial Addition). The small effect sizes for WMS−IV indexes and subtests indicate little clinical relevance for the differences since there is considerable performance overlap between men and women. Both the WAIS−IV and WMS−IV show only small differences between men and women, therefore, there are no gender-specific norms. However, sex based normative adjustments are provided in the ACS for both the WAIS−IV and WMS−IV as part of fully adjusted normative scores.

Education Effects

The effects of education on cognitive test performances have been well studied. Education has a complex relationship with cognitive test performance, particularly in clinical populations. The relationship between education and performance is likely bi-directional with good cognitive ability likely impacting educational attainment and educational attainment improving performance on cognitive instruments. The influence of education is highly variable across abilities, creating different profiles across educational levels. The expression of clinical disorders, particularly dementia, is also likely to be impacted by educational attainment. Each of these relations will be discussed in more depth.

Research evaluating the effects of education on cognitive test performance identifies significant advantages to individuals with greater educational attainment. The relationship between educational attainment and cognitive test performance is complex. Individuals with greater educational attainment are exposed to more test taking situations, may learn better test taking skills, and are exposed to more information than those with less education. These factors provide a test-taking and content exposure advantage (e.g., vocabulary) as a function of education level. Therefore, education provides an advantage to test-takers that may not necessarily relate to better cognitive abilities but rather is the outcome of greater educational experiences.

Given that general ability influences both educational attainment and test performance, individuals with high cognitive ability will perform

well on cognitive tests and will be likely to obtain more education. Therefore, adjusting scores for education level will unintentionally partial out variance related to cognitive ability, not just the effects of educational background. This is an important issue. If education effects on cognitive tests *only* represent the inherent abilities of the individual and the normative data are adjusted for education level, the sensitivity of the test is reduced. However, if education effects result in better or worse test performance due to non-ability factors, then a failure to adjust normative data may mask true differences in ability on the construct in question. If a highly educated person sustains a traumatic brain injury, and age-adjusted normative data is used, it might be difficult to accurately identify and quantify the person's cognitive deficits. This is because the person's prior education could have two influences on performance: (a) many pre-injury test scores were likely above the normative mean, and (b) following injury the person might simply be better equipped to perform well in a cognitive testing situation. Conversely, if an individual with low educational attainment sustains an injury, using education adjusted scores *could* mask the severity of the deficits in cognitive functioning due to the injury since ability has an impact on educational attainment.

While the impact of education on test performance is well-documented, additional benefits of educational exposure must be considered in the context of clinical evaluations. In particular, clinicians must consider that levels of education may affect the *development* and *expression* of cognitive disorders such as dementia (Murray et al., 2011). In older adults, education correlates with cognitive performance even after controlling for other background variables (Jefferson, et al., 2011). Low levels of education may be a risk factor for the development of dementia in later life, although, this is an inconsistent finding (Sharp & Gatz, 2011). In healthy older adults, higher education is associated with increased cortical thickness which may buffer the expression of cognitive impairment in patients with dementia (Liu et al., 2012) and the positive impact of education may be greater than the negative impact of neuropathology on cognitive functioning (Murray et al., 2011). Beyond the impact on cognition, education has an overall positive impact on health and longevity that is not explained by differences in socioeconomic advantages (Baker et al., 2011). While most of the research on education effects is based on healthy elderly and dementia patients, the same principle of education affecting the development and expression of cognitive impairment likely applies to other clinical groups as well, such as traumatic brain injury. The sensitivity of a test to brain dysfunction is likely affected by education.

The relationship of education and specific cognitive measures varies considerably (Dori & Chelune, 2004; Tombaugh, Kozak, & Rees, 1999).

The degree to which educational attainment correlates with cognitive test performance is proportional to the degree to which the test correlates with general cognitive ability; tests with low correlations with general ability have lower correlations with education (Heaton et al., 2003). Since cognitive tests correlate at different levels with educational attainment, the rates at which specific cognitive profiles are observed vary considerably by education level (Dori & Chelune, 2004). For example, in individuals with higher education levels, large discrepancies between verbal scores (which tend to be high) and visual—perceptual abilities are much more common than in individuals with lower education levels (Dori & Chelune, 2004). Education level also impacts variability in performance across serial assessments in healthy adults (Pearson, 2009), although education does not reduce the rate of cognitive loss associated with dementia (Zahodne et al., 2011). In a single assessment, variability within a profile that is consistent with the examinee's education level could be mistaken as indicative of cognitive dysfunction if education is not considered. In serial assessment, controlling for education level can improve the sensitivity to changes in functioning (see Chapter 6).

High education levels can mask cognitive difficulties in patient populations, reducing the sensitivity of the tests in these populations. Similarly, low education levels can result in misidentification of cognitively healthy individuals as cognitively impaired. When education effects are not taken into consideration, over 30% of healthy adults with fewer than 12 years of education would be identified as memory impaired on the WMS—III (Heaton et al., 2003). On a battery of neuropsychological tests, healthy adults with low education levels perform in the impaired range on multiple measures, a performance that gives the appearance of global cognitive impairment (Belzunces dos Santos, de Souza Silva Tudesco, Caboclo, & Yacubian, 2011). Considering the impact of education on test performance can improve the specificity to low, but normal, cognitive functioning.

It has been argued that education level correlates with cognitive test performance in normally developing, healthy, individuals but not in clinical populations (Reitan & Wolfson, 1996). Therefore, adjusting scores for education level would only reduce the sensitivity of neuropsychological tests to brain injury. However, education is correlated with test performance in patient populations for many cognitive tests (Boone, Victor, Wen, Ranzani, & Ponton, 2007) and in some cases, education level is more highly correlated with cognitive performance in patient populations than in healthy controls (Randolph, Lansing, Ivnik, Cullum, & Hermann, 1999). Education does result in performance differences within a variety of clinical populations. Therefore, on tests where scores correlate with education, it is important for clinicians to

consider the examinees education level when determining if a patient's score appears unexpectedly low. However, it is difficult for clinicians to know the correlation between a specific test and education level without statistical data.

Table 4.3 presents WAIS–IV index and subtest data by education level. Direct comparisons between two adjacent education levels, 12 years and 13–15 years of education, and between the extremes of the distribution, 8 or fewer years and more than 18 years, are provided. The overall distribution of mean scores across education levels illustrates the relationship between education and specific cognitive skills. Given the large sample sizes, even small differences between groups will yield statistically significant results; therefore, to get a better sense of the impact of education on test performance effect sizes should be evaluated.

The range of scores varies across education levels by index and subtest. At the index level, Verbal Comprehension shows the largest score variance ranges from 81.3 to 115.5. The difference between the highest and lowest education levels produces an effect size of −2.69. This is a very large effect size and indicates that education level will have a very big impact on the expression of impairment at the extremes of the distribution. For example, someone with more than 18 years of education would need to lose over 30 points on average to obtain a score 1 standard deviation below the mean; however, the mean for individuals with 8 or fewer years of education is already below 1 standard deviation below the mean. In the middle of the distribution, the difference between a high school diploma and some college produces a moderate effect for Verbal Comprehension, −0.5. The difference between adjacent groups is less extreme and the effect of adjusting scores is much smaller than the impact observed when comparing performances at the extremes. Finally, for groups near the mean in the standardization sample, such as the 13–15 year group, education adjustments will yield no interpretable difference in performance from traditional standard scores. Therefore, education-adjusted norms will be less useful in these groups.

Of the remaining index scores, the smallest range of scores is observed on Processing Speed (89.3 to 105.7). The effect size between the extreme groups is −0.97, still large but a smaller impact on the expression of loss of cognitive functioning than VCI. Additionally, note that for the high education group, there is a 10-point discrepancy between VCI and PSI with VCI being much higher than PSI. In the lowest education sample, the opposite pattern is observed. These naturally occurring differences in cognitive profiles can cloud the interpretation of differences in index performance (i.e., large VCI vs. PSI split indicative of brain injury). For the comparison of groups in the middle of the

TABLE 4.3 WAIS—IV Core Index and Subtest Mean Scores, by Education Level

	Education Level (By Years)														Group Differences			
	8 or Less		9—11		12		13—15		16		17—18		More than 18		12 *vs.* 13—15		*Vs.* More than 18	
Score	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Effect Size**	***p***	**Effect Size**	***p***
Verbal Comprehension Index	81.3	13	86.7	13.9	95.2	12.8	101.6	12.5	108	12.8	110.9	11.8	115.5	11.3	−0.5	<0.001	− 2.69	<0.001
Perceptual Reasoning Index	86.5	12.3	89.7	13.6	97.4	14.8	100.4	14.2	104.4	15.1	106.5	14.2	106.1	14.3	−0.2	<0.01	− 1.56	<0.001
Working Memory Index	84.3	12.5	88.7	14	97	14	100.8	14.1	106	14.2	107.4	12.9	109.5	13	−0.3	<0.001	− 2.00	<0.001
Processing Speed Index	89.3	17.3	90.7	15.3	98.3	14.7	101.2	14	104	14.5	106.1	13.8	105.7	15	−0.2	<0.05	− 0.97	<0.001
Full Scale IQ	82	12.6	86.4	13.8	96.2	13.7	101.4	13.1	107.1	14	107.1	14	111.7	12.5	−0.4	<0.001	− 2.36	<0.001
Vocabulary	6.5	2.2	7.4	2.6	9.1	2.6	10.5	2.6	11.6	2.7	12.3	2.6	13.1	2.3	−0.5	<0.001	− 2.98	<0.001
Similarities	6.9	2.9	7.9	2.9	9.3	2.6	10.1	2.5	11.3	2.4	11.7	2.5	12.7	2.1	−0.3	<0.001	− 2.11	<0.001
Information	6.7	2.6	7.6	2.8	9.1	2.7	10.3	2.8	11.6	2.8	12	2.6	12.6	2.6	−0.4	<0.001	− 2.26	<0.001
Block Design	7.9	2.9	8.4	2.8	9.6	3	10.1	2.9	10.8	3.1	11.1	2.9	10.6	3.1	−0.1	>0.05	− 0.94	<0.001
Matrix Reasoning	7.4	2.5	8	2.9	9.5	3.1	10.2	2.9	11.1	3.1	11.6	3	12.2	3	−0.2	>0.05	− 1.86	<0.001
Visual Puzzles	7.9	2.4	8.4	2.7	9.7	3.1	10.1	3	10.6	3.2	10.9	3	10.6	2.9	−0.2	>0.05	− 1.08	<0.001
Digit Span	7.5	2.9	8.3	3.1	9.7	2.9	10.1	2.9	11.1	3	11.1	2.7	11.5	3.1	−0.2	>0.05	− 1.37	<0.001
Arithmetic	7	2.2	7.8	2.5	9.3	2.7	10.2	2.8	11.1	2.9	11.6	2.7	12	2.5	−0.3	<0.001	− 2.19	<0.001
Coding	7.7	3.2	8	3	9.6	2.9	10.3	2.9	10.8	2.9	11.4	2.7	11.6	3.2	−0.2	<0.01	− 1.23	<0.001
Symbol Search	8.3	3.8	8.6	3.3	9.7	3	10.2	2.9	10.7	3.1	10.9	2.9	10.9	2.9	−0.1	>0.05	− 0.69	<0.001

Note: Sample size 0—8 years = 220, 9—11 years = 243, 12 years = 647, 13—15 years = 553, 16 years = 267, 17—18 years = 297, more than 18 years = 43.

distribution, the effect size for PSI is only 0.20 suggesting very small performance changes between adjacent groups.

Another important characteristic of education effects on cognitive variables is that the relationship is not typically linear across education levels. In particular, observe that score change is smaller for most measures from 16 years to greater than 18 years than between other points in the distribution. A linear model would require the scores to be higher than they actually are at the higher education levels. Similarly, at the lower end of the distribution, a linear model would require lower scores than were actually obtained. In addition to score change, the standard deviation of scores is not necessarily the same across education groups. For VCI, scores become less variable as education increases. What is not illustrated in Table 4.3 is the skew of the distribution that can also change in meaningful ways across education level (e.g., from positive to negative skew). Traditional regression models, even non-linear models, do not control for changes in variability or skew across the regression line. To properly model education effect, novel norming techniques are required.

At the subtest level, Vocabulary shows the largest range between the lowest and highest educational groups, more than any other variable on the WAIS–IV ($d = -2.98$). Clearly, education effects play an important role on this test. Also, individuals of high ability tend to score highest on these tests; therefore, they will need to lose a great deal of function in this skill to show a deficit. Vocabulary, and VCI more generally, are considered resistant to decline; therefore, a patient needs to lose a lot more function on these tests in order to be considered deficient. In other words, the rate of cognitive decline is slower on these tasks early in a progressive decline and scores tend to be higher than those on other skills in highly educated patients. Therefore, Vocabulary and VCI will remain higher than other cognitive skills. Adjusting for education will have the most profound effect on these scores. On the other hand, Symbol Search shows the smallest differences of any score on the WAIS–IV. Therefore, education adjustments will have only a small effect on observed results for Symbol Search.

Table 4.4 presents descriptive statistics for the WMS–IV, by education level. Among WMS–IV Indexes, the Visual Working Memory index has the largest difference between the highest and lowest education groups ($d = -1.67$). While this difference is not as large as seen for Verbal Comprehension, it is similar to the WAIS–IV Perceptual Reasoning Index. Education shows the smallest effect on the WMS–IV Delayed Memory Index ($d = -0.89$) which is similar to the WAIS–IV Processing Speed Index effect size. At the subtest level, the largest education effect is observed on Spatial Addition ($d = -1.47$) and the smallest effect size is on Visual Reproduction II ($d = -0.49$). Adjacent education groups show small differences with the biggest differences

TABLE 4.4 WMS—IV Core Index and Subtest Mean Scores, by Education Level

	Education Level (By Years)														Group Differences			
	8 or Less		9–11		12		13–15		16		17–18		More than 18		12 *vs.* 13–15		8 or Less *vs.* More than 18	
Score	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Effect Size	*p*	Effect Size	*p*
Immediate Memory Index	88.8	15.2	90.8	14.6	97.5	14.1	100.9	15.2	103.7	14.1	106.5	13.5	109.7	10.1	−0.27	<0.05	−1.45	<0.001
Delayed Memory Index	91.7	17.5	91.8	14.7	98	14.1	100.1	14.9	103.2	14.1	106.2	14.4	106.6	12.4	−0.15	>0.05	−0.89	<0.001
Auditory Memory Index	91.5	17	92.5	14	97.7	14.7	100.5	14.4	103.1	13.9	106.1	14.9	107.1	12.7	−0.20	>0.05	−0.96	<0.001
Visual Memory Index	91.2	15.9	92.2	15.2	98.3	14.7	100.2	14.7	102.8	14.7	105.4	12.9	106.5	13.2	−0.14	>0.05	−0.99	<0.001
Visual Working Memory Index	85.2	13.5	86.9	14.7	97.4	13.7	99.6	15	103.8	14	107.9	12.5	108.6	17.2	−0.14	>0.05	−1.67	<0.001
Logical Memory I	8.2	3.4	8.5	3	9.6	2.9	10.1	2.8	10.6	2.8	11.1	2.9	11.1	2.5	−0.18	>0.05	−0.90	<0.001
Logical Memory II	8.7	3.7	8.8	2.9	9.6	3.1	10.1	2.9	10.3	2.8	11	3	10.9	2.6	− 0.16	> 0.05	−0.62	<0.01

(*Continued*)

TABLE 4.4 (Continued)

	Education Level (By Years)														Group Differences			
	8 or Less		9–11		12		13–15		16		17–18		More than 18		12 *vs.* 13–15		8 or Less *vs.* More than 18	
Score	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Effect Size**	***p***	**Effect Size**	***p***
Verbal Paired Associates I	8.7	3	8.9	2.7	9.6	2.9	10.1	3	10.6	3	11	3.1	11.6	3	−0.19	>0.05	−0.96	<0.001
Verbal Paired Associates II	8.6	3.1	8.8	2.9	9.6	3	10	3	10.7	2.9	11	3	11.2	2.8	−0.12	>0.05	−0.84	<0.01
Visual Reproduction I	8.3	3.3	8.6	3.1	9.6	3.1	10.1	3	10.7	3.1	11	2.4	11.4	2.3	−0.16	>0.05	−1.00	<0.001
Visual Reproduction II	9	3.6	8.7	3.1	9.7	2.9	10	2.9	10.5	3	10.8	2.8	10.7	2.7	−0.11	>0.05	−0.49	>0.05
Designs I	7.5	2.5	8.1	2.8	9.8	2.9	10.2	3.1	10.5	2.7	10.7	2.9	10.8	2.9	−0.14	>0.05	−1.25	<0.01
Designs II	8.5	1.9	8.4	2.8	9.9	2.9	9.9	2.9	10.3	2.9	11	3.2	10.8	3.4	0.01	>0.05	−1.02	>0.05
Spatial Addition	7.2	2.4	7.9	2.7	9.5	2.8	10	3	10.8	2.8	11.6	2.7	11.2	3.8	−0.14	>0.05	−1.47	<0.001
Symbol Span	8.1	3.1	8.2	3	9.7	2.8	10	2.9	10.6	2.9	11.1	2.8	11.7	2.9	−0.13	>0.05	−1.16	<0.001

Note: Sample size 0–8 years = 149 (69), 9–11 years = 167 (91), 12 years = 438 (247), 13–15 years = 348 (238), 16 years = 172 (123), 17–18 years = 215 (142), more than 18 years = 30 (16), spatial addition and design memory in parentheses.

occurring between 9–11 years versus 12 years of education, the same groups as observed on the WAIS–IV.

The findings for the education effects on the WAIS–IV and WMS–IV suggest that high levels of education may mask acquired impairments in cognitive functioning because patients need to lose a lot of ability before their performance reaches an impaired range. Similarly, those with very low levels of education may appear to be impaired without having lost any functioning. When age-adjusted standard scores are used, patients may show deficits in domains such as memory and processing speed, since these are less impacted by education and are susceptible to the effects of brain injury, but appear to have relatively intact verbal, perceptual, and working memory abilities, tasks related to education. If education adjusted norms are applied, the sensitivity and specificity should improve for patients at the ends of the distribution but have less impact on patients with average levels of educational attainment.

Tables 4.5 and 4.6 contain subtest scaled score to T-score conversions for Vocabulary and Visual Reproduction I, respectively. These tables are used by the ACS scoring assistant to convert subtest scaled scores to education-adjusted T-scores and illustrate the degree that scores are adjusted for education level. For Vocabulary, large adjustments in relative performance can be observed at the extremes of the distribution. Examinees with low education levels obtaining low average scaled scores of 6 and 7, obtain education adjusted scores in the average range. In individuals with more than 18 years of education, average scores of 8–10 are adjusted to be in the mild/moderate to mild impairment range. Visual Reproduction, however, yields much less dramatic adjustments even at the extremes of the distribution. Scaled scores of 10 remain average across all education levels. Scaled scores of 8 become low average for examinees with more than 16 years of education. For individuals with an average level of education (i.e., 12, 13–15 years), scores are not adjusted dramatically and for most variables do not change the interpretation of the scaled score.

Of all background characteristics, education is the most strongly associated with test performance. Specifically, individuals at the extreme ends of the distribution may obtain low or high scores due to education effects. These effects may mask deficits or create an impression of impairment in healthy individuals.

Race/Ethnicity Effects

Racial/ethnic differences in test performance occur on cognitive tests in healthy control (Manly, 2005) and clinical populations (Boone et al.,

TABLE 4.5 Scaled Score to Education Adjusted T-Score Conversion Table for Vocabulary

	Education Level (In Years)						
VC SS	**0–8**	**9–11**	**12**	**13–15**	**16**	**17–18**	**>18**
1	18	15	10	10	10	10	10
2	26	22	16	13	10	10	10
3	32	28	22	18	14	10	10
4	38	35	28	23	19	15	10
5	43	40	33	28	24	20	15
6	49	45	38	33	28	25	20
7	53	49	42	37	32	29	25
8	57	53	47	41	37	34	30
9	61	56	50	45	41	38	34
10	65	60	54	49	45	41	38
11	69	63	57	52	48	45	42
12	73	66	61	56	51	49	46
13	77	69	65	60	55	53	50
14	80	73	69	64	59	57	54
15	84	77	73	67	62	61	57
16	88	81	77	71	66	64	61
17	90	85	81	75	70	68	65
18	90	90	85	79	74	72	69
19	90	90	89	83	79	76	74

2007). When considering the differences between racial/ethnic groups, there are a number of issues to consider. A detailed discussion of all the issues related to racial/ethnic group differences is beyond the scope of this chapter. Comprehensive reviews are provided in Weiss et al. (2006) and Weiss, Chen, Harris, Holdnack, & Saklofske (2010). Although not all issues are discussed here, some important aspects will be addressed.

A common misperception suggests that cognitive tests contain significant item bias that produces the differences observed between ethnic/racial groups. However, the application of modern statistical methods is used to identify and eliminate such items. Demographics, such as race,

TABLE 4.6 Scaled Score to Education Adjusted T-Score Conversion Table for Visual Reproduction I

	Education Level (In Years)						
VR II SS	**0–8**	**9–11**	**12**	**13–15**	**16**	**17–18**	**>18**
1	26	22	21	21	20	19	16
2	29	26	24	24	23	21	19
3	32	29	27	27	26	24	22
4	35	33	30	30	28	26	25
5	38	36	33	33	31	30	29
6	41	40	37	36	34	33	32
7	44	44	40	40	38	37	36
8	47	47	44	43	41	40	40
9	50	50	48	47	44	44	44
10	54	54	51	50	48	48	48
11	57	57	55	54	52	52	52
12	59	59	59	57	56	56	56
13	62	62	62	60	59	59	59
14	65	65	65	63	62	62	62
15	68	68	68	65	64	64	64
16	70	70	70	68	67	67	67
17	73	73	73	71	70	70	70
18	76	76	76	74	73	73	73
19	79	79	79	77	76	76	76

ethnicity, and educational level, are proxy variables for a host of interacting and potentially additive environmental factors that more directly influence the development, maintenance and decline of cognitive abilities across the life span. Ethnic/racial differences can be diminished by controlling for some of these environmental factors.

Cognition is malleable, to an extent, by environmental factors that mediate opportunities for cognitive growth and maintenance of cognitive abilities and the effects of these factors may be cumulative across the life span. Enriching, stimulating environments enhance cognitive development and the maintenance of cognitive abilities; whereas

impoverishing environments inhibit that growth. Further, the factors that inhibit cognitive enrichment interact with each other, such that the presence of one factor makes the occurrence of other inhibitory factors more probable. The negative effects of cognitively impoverishing environments interact over the course of a lifetime such that the impact worsens with age. At the same time, historical patterns of immigration and racism limit opportunities for quality education, occupational advancement, and access to quality health care, thereby impoverishing the cognitive developmental trajectories of generations of cultural and linguistic minority groups. When the background environmental factors are considered, racial/ethnic differences are significantly reduced.

For adults ages 20 to 90 administered the WAIS–IV, race or ethnicity explains 15 and 11% of the variance in FSIQ score differences in the African-American/White (AA/W) and Hispanic/White (H/W) comparisons, respectively. In contrast, education, occupation, income, region and gender account for 35.1% of the variance in these score differences, respectively. After controlling for this second set of variables, race or ethnicity explain 9.2 and 3.8% of the variance in AA/W and H/W score differences, respectively.

For adolescents ages 16 to 19 administered the WAIS–IV, race or ethnicity explains 5.4 and 12% of the variance in FSIQ score differences in the AA/W and H/W comparisons, respectively. In contrast, parent education, occupation, income, region, and gender account for 22.5 and 29.3% of the variance, respectively. After controlling for this second set of variables, race or ethnicity explain only 1.5 and 1.9% of the variance in AA/W and H/W score differences, respectively.

As elaborated by Weiss and colleagues (2006, 2010), data supports the thesis that observed score differences among demographic groups are not due to race or ethnicity, but to powerful environmental factors that underlie those differences. One would then expect the differences to diminish as the effects of historical racism slowly abate over generations, and as patterns of immigration and acculturation among Hispanics shift over time. Weiss et al. (2010) examined this issue by segmenting the WAIS–IV standardization sample into five birth cohorts and testing the demographic differences. There was a strong trend toward decreasing score differences with younger cohorts. As shown in Table 4.7, the score differences are reduced for younger ages. Racial and ethnic differences in IQ have been reduced by more than half a standard deviation between 1917 and 1991, a reduction only partially explained by increases in educational attainment across the generations.

Cognitive test performance relates to a number of different socioeconomic status (SES) factors, such as rates of premature births or prenatal health care, exposure to violent crime, and quality of education (McDaniel, 2006). Parent occupation and education level, and early

TABLE 4.7 WAIS—IV Mean FSIQ Difference Scores between Racial/Ethnic Groups, by Birth Cohort

	Birth Cohort				
Groups	**1917—1942 (Ages 65—90)**	**1943—1962 (Ages 45—64)**	**1963—1977 (Ages 30—44)**	**1978—1987 (Ages 20—29)**	**1988—1991 (Ages 16—19)**
AA/White	19.3	17.2	13.1	13.4	10
Hispanic/ White	17.9	13.62	14.2	7.3	9.3

Note: Ages shown are at time of standardization testing in 2007.

educational experiences also relate to cognitive test scores and vary among ethnic groups (Byrd et al., 2006). Historically, significant monetary differences invested in the education of minority children versus white children significantly impacted the educational experiences of older African-Americans (Lucas et al., 2005). Therefore, the attained education level in older African-Americans may not be equivalent to the educational experiences of whites at the same level. However, African-Americans are a heterogeneous group and the degree to which they have experienced economic, health, occupational, or educational disadvantage varies. While it is important to consider racial/ethnic differences when interpreting test scores, considering the background of the individual is also important in making appropriate clinical decisions.

A failure to consider racial/ethnic group differences can result in misinterpretation of obtained test scores. Specifically, misinterpretations can occur when applying age only adjusted norms to identify cognitive deficits indicative of acquired brain injury or cognitive decline. When standard norms are applied, 15 to 20% of healthy Hispanics and up to 35% of healthy African-Americans may be misidentified as having general cognitive or memory dysfunction compared to 10—14% of Whites (Heaton et al., 2003). Tables 4.8 and 4.9 present WAIS—IV and WMS—IV descriptive data by racial/ethnic groups. In addition, comparison data is provided for White *vs*. African-American differences. The effect sizes for WAIS—IV indexes range from 0.8 (WMI, PSI) to 1.1 (PRI, FSIQ) and from 0.6 (Digit Span) to 1.1 (Block Design) for WAIS—IV subtests. The effect sizes for WMS—IV indexes range from 0.6 (AMI, VMI, DMI) to 0.7 (IMI, VWMI) and from 0.4 (VPA I and II, VR II) to 0.7 (LM I and II, SA, SY) for subtests.

In general, WAIS—IV shows large effect sizes and WMS—IV shows moderate differences between Whites and African-Americans.

TABLE 4.8 WAIS–IV Core Index and Subtest Mean Scores, by Ethnicity

	White		African-American		Hispanic		Asian		White *vs.* African-American	
Score	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Effect Size**	***p***
Verbal Comprehension Index	102.9	14.2	90.5	15	90.9	15.4	103.6	15.5	0.9	<0.001
Perceptual Reasoning Index	103	14.5	87.5	13	93.7	14.1	104.5	15.4	1.1	<0.001
Working Memory Index	102.9	14.1	91.5	14.7	90.2	14.8	104.6	15.2	0.8	<0.001
Processing Speed Index	102.4	14.5	90.6	15.5	95.7	15.1	107.4	16.1	0.8	<0.001
Full Scale IQ	103.4	14	87.7	14.4	91.1	14.5	106.1	15.5	1.1	<0.001
Vocabulary	10.6	2.8	8.4	3.1	8.2	3.1	11	3.6	0.8	<0.001
Similarities	10.5	2.7	8.4	2.9	8.4	2.9	10.4	2.7	0.8	<0.001
Information	10.5	3	8.1	2.9	8.5	3.2	10.7	2.9	0.8	<0.001
Block Design	10.6	2.9	7.6	2.6	9.1	2.8	11	3.3	1.1	<0.001
Matrix Reasoning	10.6	3	8.2	3.1	8.7	3.1	11	2.8	0.8	<0.001
Visual Puzzles	10.5	3.1	7.9	2.4	9	2.8	10.3	2.8	0.9	<0.001
Digit Span	10.5	2.9	8.8	3.1	8.3	3.1	10.8	3	0.6	<0.001
Arithmetic	10.6	2.9	8.2	2.7	8.3	2.7	11	3.2	0.8	<0.001
Coding	10.5	2.9	8.3	3.1	9	3	11.7	3.5	0.7	<0.001
Symbol Span	10.4	3	8.3	3.2	9.4	3.3	11.1	3.3	0.7	<0.001

Note: Sample Size White = 1355, African-American = 400, Hispanic = 399, Asian = 60.

Subsequently, a failure to account for ethnic group differences could mask cognitive symptoms in whites or suggest cognitive impairment in healthy African-Americans.

Next to education level, race/ethnicity has the most impact on observed test performance. Studies indicate that Hispanics and African-Americans are at higher risk for misidentification of cognitive impairment in the absence of a disease process. The average performance of Whites and Asians is slightly higher than the overall population that potentially can affect the sensitivity of the WAIS–IV and

TABLE 4.9 WMS—IV Core Index and Subtest Scores, by Ethnicity

	White		African-American		Hispanic		Asian		White *vs.* African-American	
Score	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**ES**	***p***
Immediate Memory	101.9	14.9	91.3	14.7	94.6	14.6	104.4	14	0.7	<0.001
Delayed Memory	101.4	14.8	92.1	15.4	96.4	15.2	105	13.8	0.6	<0.001
Auditory Memory	101.9	14.9	92.5	14.8	96.1	15.5	101.1	13.8	0.6	<0.001
Visual Memory	101	14.7	92.5	15.2	96.2	14.3	107.9	14.7	0.6	<0.001
Visual Working Memory Index	102.1	14.6	91	15.5	95.6	14.9	108.2	15.8	0.7	<0.001
Logical Memory I	10.4	2.9	8.5	2.9	9	3.3	9.9	2.7	0.7	<0.001
Logical Memory II	10.4	3	8.5	2.9	9.3	3.3	9.9	2.7	0.7	<0.001
Verbal Paired Associates I	10.2	3	9	2.9	9.6	3	10.4	2.8	0.4	<0.001
Verbal Paired Associates II	10.2	3	9	3.3	9.5	3	10.6	2.9	0.4	<0.001
Visual Reproduction I	10.3	3	8.6	3.1	9.1	3	11.2	2.6	0.6	<0.001
Visual Reproduction II	10.1	3	9	3	9.6	3.2	10.8	2.7	0.4	<0.001
Designs I	10.3	3	8.7	2.9	9.2	2.9	11.7	2.7	0.5	<0.001
Designs II	10.2	3	8.9	2.6	9.5	2.7	12.1	3.3	0.5	<0.001
Spatial Addition	10.3	3	8.4	2.9	9.5	3	11.8	3.2	0.7	<0.001
Symbol Span	10.4	2.9	8.4	2.9	8.9	3.1	10.5	3.1	0.7	<0.001

Note: Sample Size White = 933 (503), African-American = 259 (174), Hispanic = 264 (206), Asian = 39 (30), spatial addition and designs in parentheses.

WMS—IV in identifying cognitive deficits in these groups. Normative adjustments using race/ethnicity may improve the sensitivity and specificity of the tests in clinical application.

DEMOGRAPHIC ADJUSTMENTS TO NORMS

Given the extensive body of research demonstrating differences on cognitive tests by education, sex, and ethnicity, neuropsychological studies have proposed developing alternate norms for subgroups of

examinees. Karzmark, Heaton, Grant, and Matthews (1984) applied multiple regression techniques to adjust normative data on the Halstead-Reitan Neuropsychological Battery Average Impairment Index to account for the significant correlations with education, sex, and ethnicity. Additionally, educational adjustments exist for Trail Making Tests A and B (Tombaugh, 2004), CVLT (Norman, Evans, Miller, & Heaton, 2000), PASAT (Gonzalez, 2006), and the Boston Naming Test (Heaton, Avitable, Grant, & Matthews, 1999). Ethnicity adjustments exist for the Stroop test (Moering, Schinka, Mortimer, & Graves, 2004), Symbol Digit Modalities Test (Gonzalez et al., 2007), CVLT (Norman et al., 2000), WAIS–R/WMS–R (Lucas et al., 2005), and the PASAT (Gonzalez et al., 2006). Ethnicity adjusted normative data improves diagnosis of dementia among older African Americans (Lucas et al., 2005). Full demographic corrections (i.e., education, gender, and ethnicity) were developed for the WAIS–R (Heaton et al., 1991) and the WAIS–III/WMS–III (Taylor & Heaton, 2001). The development and use of demographic adjustments has been researched for several decades.

Even though research has demonstrated the utility of demographic norms, no consensus exists about the clinical application of demographic adjustments. Reitan and Wolfson (1996) argue that it is not appropriate to make demographic adjustments in the evaluation of patients with brain injury because it makes the test less sensitive to the effects of the injury. Unadjusted raw scores (e.g., not even age adjustments) show greater cognitive impairment in clinical populations compared to using demographic adjustments (Golden & van den Broek, 1998). Both, adjusted and unadjusted scores can equally identify focal lesions (Golden & van den Broek, 1998). Alternatively, Bernard (1989) found that unadjusted scores over-classify healthy Whites with low education levels, Hispanics, and African-Americans as cognitively impaired. The neuropsychological evaluation of educationally and ethnically diverse populations is a complex issue (Brickman, Cabo, & Manly, 2006).

There are clear advantages to using demographic adjustments to normative data. When evaluating individuals with low educational attainment or ethnic/racial minorities, using only age-adjusted or raw scores decreases the specificity of cognitive tests in identifying brain injury or dysfunction (Heaton et al., 1999, 2003; Norman et al., 2000). Also, when evaluating individuals with high levels of educational attainment, demographically adjusted norms improve diagnostic sensitivity (Morgan et al., 2008). While, demographic adjustments to norms show good sensitivity and specificity for a number of clinical populations (Taylor & Heaton, 2001), their primary advantage is improving diagnostic specificity for minorities (Manly, 2005) and examinees with low education levels (Heaton et al., 1999).

Although there are advantages to adjusting norms for ethnic group differences, there are also disadvantages. Adjusted scores may be misunderstood or inappropriately applied. There is no scientific method for identifying an individual's race or ethnicity (e.g., these are primarily social/political constructs) and the background/social forces for which race/ethnicity serve as a "proxy" are not completely understood (Manly, 2005). Using race-based adjustments can result in examinees not getting needed services because their adjusted scores are higher and may fall above an established cut-off (Manly & Echemendia, 2007). Applying adjustments for racial/ethnic group membership, fails to account for the underlying cultural, health, and educational factors that produce disparities in test performance. Therefore, the adjustments do not necessarily represent characteristics of the individual but rather represent general characteristics of the group (Manly, 2005; Manly & Echemendia, 2007). It is important to consider that within groups (e.g., ethnic, education, or sex) variability is greater than between group variability. The clinician must be clear about the rationale for using demographic adjustments to norms and the impact that such adjustments can have on the outcomes of an evaluation.

When to Use Demographic Adjustments

As a clinician, the decision about whether or not to use demographically adjusted norms depends primarily on the decisions or conclusions that need to be made. For example, if the purpose of the evaluation is to determine the examinees functioning relative to the general population (e.g., ability to work in any capacity, to live independently, or to understand social conventions; capacity to consent to treatment, etc.), then using standard age-adjusted norms is appropriate. If the evaluation requires a decision regarding the examinees ability to function in a specific environment (e.g., "*Can the examinee return to his job as a physician?*") then demographic adjustments may be appropriate. If the primary purpose of the evaluation is to identify cognitive deficits that may indicate brain dysfunction or decline in cognitive functioning, using demographic adjustments is appropriate. However, it would be inappropriate to use demographic norms to diagnose mental retardation or learning disability because these disorders are diagnosed relative to the general population (e.g., someone is not "learning disabled" compared to other Caucasians or compared to other people with Master's Degrees).

In order to appropriately use demographic adjustments, the clinician should understand how the adjustments change the scores. In other words, how does the examinee's relative rank order on the test change

with adjustments and what are the consequences of reporting lower or higher scores for that individual. Note that because scores change metrics from standard scaled scores to T-scores, it is important for comparative purposes to evaluate the percentile rank changes after demographic adjustments. Adjustments made near the middle of the distribution (e.g., education levels 12, 13–15) result in very small rank order changes, particularly on memory measures. Substantially larger effects are observed at the extremes, which include Asians and Whites with high education levels, and Hispanics and African-Americans with low education levels. For high ability groups, demographic adjustments always return a lower score (i.e., percentile rank and distance from the mean) compared to the standard age-adjusted norms. For lower ability groups, the adjusted scores will always be higher (i.e., percentile rank and distance from the mean) than the standard age-adjusted scores.

If having a low score results in some benefit (e.g., eligibility for some service, or receiving medical treatment), the potential exists that using demographic adjustments may deny those benefits to the examinee. This potential negative consequence should not completely influence the decision to use demographic adjustments; however, the clinician needs to be very clear about the appropriateness and rationale for their use. Similarly, adjusting a score higher can also have a negative consequence other than denial of a benefit. For example, in death penalty cases, having a higher score can influence the decision to pursue the death penalty instead of a lesser sentence, a great consequence to the individual. Clearly, the onus is on the clinician to demonstrate that comparing the examinee to individuals of a similar background is appropriate. In other words, is an examinee with historically low cognitive ability compared to the general population (e.g., low educational attainment, low IQ scores on testing) more culpable than someone with higher education who has an average IQ compared to the general population but whose demographic adjusted scores are in the low range of cognitive ability. The consequences for the individual and for society in general must be considered. If the purpose of the evaluation is to answer a question about a *change* in cognitive functioning, then the use of demographic norms is easily defended, if not then the examiner needs to have a sound rationale for adjusting the normative data.

Development of Demographic Adjustment to Norms

Many statistical models have been applied to developing adjustments to normative data. Primarily, regression based models are used (Karzmark et al., 1984; Heaton et al., 2003). There are clear advantages to using regression techniques, as long as the correct model is applied

and it is consistent with the underlying assumptions of the relationship between the background and cognitive variables being adjusted. The primary assumption made during the development of the WAIS−IV and WMS−IV demographic adjustments was that higher education is associated with better or equal cognitive skills than lower levels of education, across all groups. In other words, there is no theoretical reason to assume that African-American males with a high level of education have *worse* cognitive skills than African-American males with an average degree of education. While the assumption may seem obvious, sampling artifacts can produce unexpected findings and some regression techniques will follow the sample data, yielding unexpected results. Additionally, the assumptions of equal variances (i.e., homoscedasticity) and equal skew across the regression line were assumed to not always hold true. When these conditions are false, the regression method systematically makes errors in the adjustment of scores along the regression line. Therefore, inferential norming techniques were applied to the derivation of the WAIS−IV and WMS−IV demographic adjustments.

The demographic adjustments are derived from the WAIS−IV and WMS−IV normative samples and an additional oversample of individuals with low and high education levels, and minorities. The oversampling is important because the standardization sample may not include sufficient cases for estimating the ability of certain groups (e.g., Hispanics with 17−18 years of education).

All interactions were evaluated (e.g., age by education, education by sex, education by sex by ethnicity) to determine the most appropriate model for adjusting the normative data. Given that education only norms were required as part of the WAIS−IV/WMS−IV project, education adjustments were developed prior to adjusting for sex and ethnicity. Education level is segmented into seven bands: 8 years or less, 9 to 11 years, 12 years, 13 to 15 years, 16 years, 17 to 18 years, and over 18 years. For full demographic adjustments, the sample is divided into groups as follows: White male, White female, African-American male, African-American female, Hispanic male, Hispanic female, Asian-male and Asian female. These samples are not stratified by census but include higher percentages of examinees from lower and higher education groups and more ethnic minorities.

Derivation of Education Subtest Adjusted T-Scores

The WAIS−IV and WMS−IV demographic adjustments are derived using inferential norming (Gorsuch, 2003; Roid, 2003; Wilkins, Rolfhus, Weiss, & Zhu, 2005). Various moments (means [M], standard deviations [SD], and skewness) of each age-adjusted subtest scaled score were calculated for each education group. The moments are plotted across education level, and various polynomial regressions ranging from linear to

4th degree polynomials were fit to the moment data. Functions for each score moment were selected based on consistency with underlying theoretical expectations and the pattern of improved cognitive skills with education level observed in the sample. For each subtest, the functions are used to derive estimates of the population moments. The estimated moments are used to generate theoretical distributions for each of the reported normative education groups, yielding mid-point percentiles for each age-adjusted scaled score within education group. The mid-point percentiles are converted to T-scores, using z-normalization, with a mean of 50, a standard deviation of 10, and a range of 10–90. The progression of standard scores within and across education groups is examined, and minor irregularities are eliminated by smoothing.

Derivation of Index Scores

The subtest adjusted T-scores were summed for each of the WAIS–IV and WMS–IV index scores. The data on these sums of T-scores demonstrated a high degree of similarity across education within each of the indexes. An analysis of variance revealed no statistically significant variation by education group in the mean scores for indexes. Consequently, the education groups were combined to construct the tables of index T-score equivalents of sums of subtest T-scores.

For each scale, the distribution of the sum of scaled scores was converted to a scale with a mean of 50 and a standard deviation of 10. This conversion was accomplished by preparing a cumulative frequency distribution of the actual sum of subtest T-scores, calculating mid-point percentiles, and for each index and using z-normalization calculating the appropriate index score equivalent for the sum of subtest T-scores. Successive adjustments were based on visual inspection of the distributions while attempting to keep the means and standard deviations of the scales close to 50 and 10.

Derivation of Full Demographic (Education, Sex, Ethnicity) Subtest and Index Adjusted T-Scores

Inferential norming was also used to derive the full-demographic adjustments. For full-demographic adjustments, the education adjusted T-scores for each subtest are used to derive the moments for each normative group (e.g., male vs. female). Regression functions for each score moment are selected based on consistency with observed associations between the demographic variables and the subtest scores, the pattern of curves observed in the normative sample, and the underlying theoretical model for education effects. For each subtest, the functions were used to derive estimates of the population moments. The estimated moments were then used to generate theoretical distributions for each of the reported normative sex by education groups, yielding percentiles

for each sex by education-adjusted T-score. These percentiles were converted to T-scores with a mean of 50, a standard deviation of 10, and a range of 10−90. The progression of standard scores within and across sex by education groups was then examined, and minor irregularities were eliminated by smoothing. The first step of creating sex by education norms accounts for the interaction between sex and education. The final full-demographic adjustments are made using the education by sex adjusted scores.

Functions for the four ethnic groups are derived using the sex by education-adjusted subtest T-scores. Functions for each score moment are selected based on consistency with observed associations between the demographic variables and the subtest scores, the pattern of curves observed in the normative sample, and the theoretical model of education effects. Scores are converted to fully adjusted T-scores using the procedure described previously. Finally, an additional evaluation of the distributions by education, sex, and ethnicity was done to ensure any interactions were accounted for in the final norms. Index scores are derived using the same procedure applied for the derivation of education index scores.

Table 4.10 presents full demographic-adjusted WAIS−IV Index scores by racial/ethnic group. Application of full-demographic adjustments yields no differences between the groups. The use of full demographic adjustments will increase the rank order of the scores in African-American and Hispanic groups to be equivalent to Whites and Asians. This does not necessarily make the scores more or less appropriate than education only corrections; however, the specificity of the scores for identifying brain injury and loss of cognitive functioning should improve when applied to these groups.

Intercorrelation of WAIS−IV/WMS−IV Full Demographically Adjusted Scores

Adjusting scores by demographic variables changes the relationship between scores both within and between test batteries. A detailed presentation of this phenomenon is presented in the Advanced Clinical Solutions Clinical and Interpretation Manual (Pearson, 2009). For the purposes of this chapter, it is important to have a general sense of the change in the relationship between variables of interest. In general, the correlations among WAIS−IV subtests and indexes are lower after applying demographic corrections. The correlations among WAIS−IV age-adjusted verbal subtests range from 0.64 to 0.74; however, applying full demographic adjustments reduces that range to 0.50 to 0.63. The WAIS−IV Index correlations range from 0.45 to 0.86 but only from 0.29 to 0.81 for fully adjusted T-scores. Increased variability among the

TABLE 4.10 WAIS–IV Core Index and Subtest Education and Full Demographically Adjusted Scores, by Ethnicity

	White		African-American		Hispanic		Asian		White Vs. African-American	
Score	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Effect Size**	***p***
Verbal Comprehension Index	49.8	9.8	50	10.1	50	10.2	50.6	9.9	0	> 0.05
Perceptual Reasoning Index	50	10.1	50.1	10	50	10	50.4	11	0	> 0.05
Working Memory Index	49.7	9.9	50.1	9.9	49.9	10.2	51.5	9.4	0	> 0.05
Processing Speed Index	49.9	10	49.9	9.9	50.2	9.8	50.5	10.1	0	> 0.05
Full Scale IQ	49.8	9.8	50	10.3	50.1	10.2	50.9	10.2	0	> 0.05

Note: Sample Size White = 1355, African-American = 400, Hispanic = 399, Asian = 60.

WAIS–IV subtests and indexes should be expected when using demographically adjusted scores.

For the WMS–IV, the correlations are lower when demographic norms are applied although differences are smaller than those observed for the WAIS–IV. Age-adjusted verbal memory subtest correlations range from 0.37 to 0.87 and for full demographic adjustments the range is from 0.36 to 0.84. The age-adjusted index correlations range from 0.48 to 0.87 and from 0.39 to 0.85 for full demographic adjustments. Correlations between the WAIS–IV and WMS–IV are also lower when applying demographic adjustments. For age-adjusted indexes, the correlations range from 0.40 to 0.71, while the range is from 0.30 to 0.61 for full demographic adjustments. Clinicians should expect more dissociation among memory measures and between intellectual and memory functioning when using demographically adjusted norms. Greater variability does not indicate pathology but is a direct consequence of lower between test correlations (see Chapter 3, which discusses Variability in WAIS–IV/WMS–IV test scores).

The practical implication, however, is that when using demographically-adjusted norms a clinician is somewhat more likely to see greater variability between index scores. Index difference scores are often used to support an inference of acquired cognitive impairment. The clinician simply needs to be aware, however, that some of this

difference, when using demographically-adjusted norms, results from how controlling variance attributable to demographic variables reduces the correlation between indexes—thus it is in part a psychometric factor that needs to be considered.

Clinical Application

Clinical data applying demographically adjusted norms is found throughout this text. In each chapter that addresses a specific clinical sample, tables comparing patients with both matched controls and a random sample of low and high education healthy controls are presented. Using matched control studies only does not illustrate the impact of demographically adjusted norms given that both the clinical group and the matched controls have scores adjusted by a constant value; therefore, the effect sizes will be very similar between using standard norms and demographically adjusted norms. Comparing clinical groups to random cases of low and high education normal controls provides a better illustration for the clinician of how the scores will enable them to identify patients from low functioning healthy individuals or to individuals with very high levels of functioning. This model better approximates the clients often seen in practice and enables a better comparison on which to make decisions about using demographic adjustments.

Caveats in Using Demographic Adjustments to Norms

The use of demographic norms requires that certain assumptions are true. The primary assumption is that the individual's demographic background is representative of their personal experiences and is a good estimate of their pre-morbid ability. There are no hard and fast rules that can be applied to answer this question; however, there are some factors to consider. The first consideration is whether or not the examinee's education level reflects their ability or not. In cases where the examinee has a chronic medical, neurological, or psychiatric condition, their education level may have been disrupted by the condition itself. Individuals with chronic epilepsy from childhood, a history of brain tumors and treatment during childhood, or any chronic disease that may have limited their ability to attend or benefit from schooling may not have reached their full academic potential. Some psychiatric conditions, such as schizophrenia, have initial onset of symptoms in adolescence and can have a negative impact on educational attainment. Similarly, individuals with learning disabilities may never achieve academically at a level consistent with their cognitive abilities. In these and similar situations, the examiner must consider if the attained level of education is a good proxy for a patient's pre-morbid intellectual functioning.

The second consideration is whether or not to use adjustments for race/ethnicity. As stated previously, racial/ethnic status cannot be

determined by any scientific means and it may not be possible to accurately classify an individual client. The clinician will need to use the examinee's own conceptualization of their race/ethnicity. The clinician will also need to determine if the individual's background is representative of the factors that can result in cognitive differences between groups. In other words, if the examinee grew up in a wealthy neighborhood, with highly educated parents, attended good schools, and is also Hispanic, do the adjustments made to the normative data accurately reflect the individual's background? It was hypothesized that socioeconomic disadvantage, health care disparities, poor educational experiences, and other potential discriminatory factors may account for the between group differences but if none of these factors is present for a specific individual, does it make sense to adjust for ethnicity? There is no simple answer to this question and in some cases it may be yes (e.g., level of acculturation in the family generally may be low, affecting language development, effects of racism limiting opportunities) or in some cases no (e.g., background is not inconsistent with other non-minority groups). The clinician must use his or her judgment as to when it is appropriate to adjust for racial/ethnic differences.

The final consideration for using demographic norms is related to whether the application will really make a difference or not. If the examinee has 13 years of education, the normative adjustments will be negligible so does it make sense to adjust the scores or not? In some cases, the answer may still be yes, if the other background factors will have an impact (e.g., African-American, female, 13 years of education) and the question concerns a change in cognitive status. In other cases, it may not be helpful even when looking at a change in function. Despite the fact that the changes may or may not be large, the best approach to using demographic norms is to be consistent. If there is a question of cognitive change, then use either the education or demographic adjustments. If a patient's background is not consistent with some of the group factors that might result in between racial/ethnic group differences, then use education only adjustments. The clinician needs to be able to rationally defend his or her choice and to apply the correct norms and interpretation to answer the specific referral questions.

CASE STUDIES

CASE STUDY 1

APPROPRIATE USE OF DEMOGRAPHIC ADJUSTMENTS Mr. F. is a 59-year-old, Hispanic, married man, who completed 9 years of education. Mr. F. has suffered from Parkinson's disease for over 15 years

with increasing motor symptoms but relatively intact cognition. He is ambulatory with a walker as long as he takes his medication as prescribed. Mr. F. was formerly employed as an auto mechanic for over 30 years but has not been able to work the past 7–8 years. He has been able to care for himself at home; however, within the past year, his family has observed a decline in his memory functioning and reasoning skills. The family is concerned about his capacity to remain at home alone and potential safety issues. His wife is still employed and cannot be home with her husband during the day. Mr. F. has been forgetting to take his medicine despite using a pill box and his wife calling home to remind him. He has lost weight and the family fears he is not remembering to eat during the day. Mr. F. was referred for neuropsychological evaluation to determine if there has been a significant change in memory functioning and to determine if there are early signs of dementia.

The WAIS–IV and WMS–IV were administered as part of a neuropsychological battery. Mr. F. was medicated for the evaluation but still showed characteristic signs of Parkinsonism including: resting tremor, rigidity, festinating gate, and masked facies. He had obvious difficulty initiating behavior but showed no signs of impaired mental status. Tests involving motor control were performed slowly and with difficulty. On the Brief Cognitive Status Exam (BCSE), his orientation, memory, and language were within normal limits. He showed slight processing speed and cognitive control weaknesses. His overall score of 47 was average for his age and education.

Table 4.11 presents Mr. F.'s WAIS–IV age-adjusted and demographically adjusted scores. His age-adjusted FSIQ of 74 (4th percentile) is in the borderline range of functioning. His index scores indicate deficient processing speed and working memory abilities. Verbal Comprehension is low average, while Perceptual Reasoning is in the average range. Clearly, there are significant strengths and weaknesses in this profile with some scores impaired while others are intact. Using age-adjusted scores, Verbal Comprehension is significantly lower than Perceptual Reasoning, and Working Memory and Processing Speed are significantly lower than Verbal Comprehension. While some of the scores are impaired and there is variability in the profile, does this represent a decline in functioning or are results consistent with Mr. F.'s background?

When demographic adjustments are applied, Mr. F.'s FSIQ is in the low average range. His Perceptual Reasoning Index is above average and is a relative strength compared to the Verbal Comprehension Index; however, Verbal Comprehension is average for his background. Working Memory and Processing Speed Index scores are in the impaired range even after controlling for demographic characteristics.

TABLE 4.11 Mr. F.'s Age-Adjusted and Full Demographic Adjusted WAIS–IV Scores

	Age Adjusted		Full Demographic Adjusted		
Score	**Score**	**Percentile Rank**	**T-Score**	**Percentile Rank**	**Qualitative Descriptor**
Verbal Comprehension Index	83	13	49	46	Average
Perceptual Reasoning Index	96	39	55	69	Above Average
Working Memory Index	66	1	35	7	Mild Impairment
Processing Speed Index	56	0.2	31	3	Mild to Moderate Impairment
Full Scale IQ	74	4	42	21	Low Average
General Ability Index	88	21	52	58	Average
Vocabulary	6	9	49	46	Average
Similarities	3	1	35	7	Mild Impairment
Information	12	75	63	90	Above Average
Block Design	6	9	42	21	Low Average
Matrix Reasoning	8	25	53	62	Average
Visual Puzzles	14	91	68	96	Above Average
Digit Span	3	1	36	8	Mild Impairment
Arithmetic	5	5	38	12	Mild Impairment
Symbol Search	3	1	35	7	Mild Impairment
Coding	1	<1	28	1	Moderate Impairment

At the subtest level, Mr. F. shows variability in Verbal Comprehension and Perceptual Reasoning domains. In the Verbal domain, he demonstrated good recall for long-term information and average vocabulary skills; he exhibited a weakness in verbal conceptual reasoning. Perceptual Reasoning is notable for good visual construction skills and mental rotation in the absence of any motor demands. Mr. F.'s Block Design performance was clearly hampered by motor initiation and control problems. Processing speed and working memory are significant weaknesses.

TABLE 4.12 Mr. F.'s Age-Adjusted and Full Demographic Adjusted WMS—IV Scores

	Age Adjusted		Full Demographic Adjusted		
Score	Score	Percentile Rank	T-Score	Percentile Rank	Qualitative Descriptor
Auditory Memory	91	27	50	50	Average
Visual Memory	81	10	43	24	Low Average
Immediate Memory	78	7	43	24	Low Average
Delayed Memory	90	25	49	46	Average
Logical Memory I	8	25	50	50	Average
Logical Memory II	9	37	51	54	Average
Verbal Paired Associates I	10	50	54	66	Average
Verbal Paired Associates II	7	16	44	27	Low Average
Designs I	8	25	50	50	Average
Designs II	10	50	56	73	Above Average
Visual Reproduction I	1	<1	27	1	Moderate Impairment
Visual Reproduction II	8	25	46	35	Average

WMS—IV standard and full demographically adjusted index and subtest scores are displayed in Table 4.12. Mr. F.'s age-adjusted index scores indicate borderline immediate recall and low average visual memory with average scores on auditory and delayed memory. Demographic adjustments show Mr. F. has low average to average memory functioning for his background. Using demographic adjustments, visual memory is significantly lower than auditory memory but the base rate is not atypical (>25%). Immediate and delayed memory scores are not significantly different from one another.

WMS—IV verbal memory subtests are generally in the average range for immediate learning and recall. His long-term cued recall for verbal associations was in the low average range that is the borderline range for his level of immediate recall on that subtest. His delayed recognition for verbal associations was average; indicating low average cued recall (contrast score) or a retrieval problem for verbally associated material. Visual memory subtest performance was remarkable for the impact of motor issues. On Designs, the examiner assisted in placing the cards in the grid but on Visual Reproduction Mr. F. struggled to draw the designs.

All scores were average except for immediate visual reproduction; which appeared to be an issue with drawing more than recall. His raw score on delayed and copy drawings were very similar to immediate recall. His copy score was also in the deficient range. Subtest memory scores were all in the average range when demographic norms were applied except for Visual Reproduction I and Verbal Paired Associates II. None of the verbal subtests are significantly different from one another when demographic norms are applied. The Designs subtest scores are significantly better than Visual Reproduction scores; and Visual Reproduction I is lower than Visual Reproduction II. Differences in visual memory functioning relate to motor skills rather than visual memory impairment. Recognition memory scores, which are not demographically adjusted, were in the borderline to high average range. A borderline score was obtained on Designs Delayed Recognition but was more likely related to inattention than memory impairment.

These test results suggest some cognitive limitations in working memory and declarative memory. Declarative memory deficits appear to be related to retrieval rather than encoding issues. Motor impairments affect performance on many of the tests making it difficult to assess the degree to which processing speed may be impaired. On non-motor based processing speed measures (BCSE mental control and inhibition), he showed lower than expected scores but not significantly impaired scores. Some degree of cognitive slowing appears to be present. The overall clinical picture does not indicate dementia; however, cognitive deficits associated with Parkinsonism in combination with increasing age are affecting his daily functioning. The family worked with the hospital social worker to arrange day programming for Mr. F. Re-evaluation in 6–9 months is recommended to track cognitive changes.

CASE STUDY 2

INAPPROPRIATE USE OF DEMOGRAPHIC ADJUSTMENTS Ms. J. is a 22-year-old, single, African-American female, referred for a disability evaluation subsequent to chronic seizure disorder. Ms. J. had been living at home with her parents and working part-time as a cashier in a clothing store. Historically, her medication has reduced but not completely controlled her seizures. Approximately 9–10 months ago, her rate of seizures increased to 4–5 times per week, interfering with her ability to work and keeping her primarily at home. Medication changes did not improve her seizure rate. About 6 months ago, she experienced a prolonged seizure episode (status epilepticus) requiring emergency hospitalization. She was successfully treated but her seizures are still not well controlled.

Ms. J. attempted to return to work 3 months ago. However, she frequently missed time from work and her employer was not able to

accommodate her absences. She has obtained employment in other businesses but she was unable to maintain work on a regular part-time basis. Her parents report that she seems to fatigue easily and has problems with memory and concentration. Ms. J.'s neurologist recommended that she apply for social security disability, even though she would prefer to work. As part of the disability evaluation, she was referred for psychological assessment to document her current level of cognitive functioning.

Ms. J. has suffered from generalized seizure disorder since she was 11 years of age. Prior to the onset of her seizures, she was consistently a top student in her classes. After her seizures, she had difficulty keeping up in her classes due to frequent absences and medication side-effects. She did not require special education courses but accommodations were made for her medical condition. She graduated high school at the age of 19. She grew up in a wealthy suburban neighborhood. Her father has a Ph.D. in chemical engineering and founded his own business. Her mother has a master's degree in speech therapy and worked in a local hospital.

The psychologist administered the WAIS−IV and WMS−IV as part of the disability evaluation. The results of the WAIS−IV are presented in Table 4.13. Her performance on the WAIS−IV was consistently in the low average range. Her FSIQ (83) was at the 13th percentile and there were no significant differences among the WAIS−IV indexes. WAIS−IV subtests ranged from borderline (Coding = 5) to Average (Vocabulary, Information, and Visual Puzzles = 9). No scores were significantly different. The age-adjusted scores consistently place her cognitive skills in the low average range. Her basic vocabulary and general knowledge were average.

The psychologist applied demographically adjusted norms to determine if there was a change in cognitive functioning. She applied full demographic adjustments (Sex = female, Education = 12 years, Ethnicity = African-American) to the WAIS−IV data. The demographic adjusted index scores are all in the average range except for the Processing Speed Index which was low average. Processing Speed was significantly lower than her Verbal Comprehension and Perceptual Reasoning Index scores. Working Memory was significantly lower than Perceptual Reasoning, but the base rate occurs relatively commonly (BR = 28.9%).

The WAIS−IV subtests ranged from mild impairment (Coding = 35) to above average (Information, Visual Puzzles = 57). In the verbal domains, Similarities was significantly lower than Vocabulary and Information. Perceptual Reasoning subtests show significantly lower Matrix Reasoning compared to Visual Puzzles. No differences occurred on Working Memory measures. For Processing Speed, Symbol Search was significantly higher than Coding.

TABLE 4.13 Ms. J.'s Age-Adjusted and Full Demographic Adjusted WAIS–IV Scores

	Age Adjusted		Full Demographic Adjusted		
Score	**Score**	**Percentile Rank**	**T-Score**	**Percentile Rank**	**Qualitative Descriptor**
Verbal Comprehension Index	89	23	51	54	Average
Perceptual Reasoning Index	88	21	53	62	Average
Working Memory Index	86	18	47	38	Average
Processing Speed Index	81	10	40	16	Low Average
Full Scale IQ	83	13	48	42	Average
General Ability Index	87	19	52	58	Average
Vocabulary	9	37	54	66	Average
Similarities	6	9	42	21	Low Average
Information	9	37	57	76	Above Average
Block Design	8	25	53	62	Average
Matrix Reasoning	7	16	47	38	Average
Visual Puzzles	9	37	57	76	Above Average
Digit Span	7	16	43	24	Low Average
Arithmetic	8	25	51	54	Average
Coding	5	5	35	7	Mild Impairment
Symbol Search	8	25	47	38	Average

The WMS–IV data is presented in Table 4.14. The age-adjusted scores range from borderline to low average. Delayed Memory is higher than Immediate Memory (contrast score = 13). Visual Working Memory is consistent with general memory functioning. WMS–IV subtest scores range from the deficient to average range. Her performance is consistent across verbal memory subtests. On visual memory, she exhibited a weakness on Visual Reproduction I and a significant strength on Designs II. Using the WAIS–IV GAI (87), all memory scores are significantly lower than general cognitive functioning (using either predicted difference or contrast score method). Most of the comparisons are in the low average range (e.g., contrast score = 7); however, immediate memory is in the borderline range given her general ability (contrast score = 5).

TABLE 4.14 Ms. J.'s Age-Adjusted and Full Demographic Adjusted WMS—IV Scores

	Age Adjusted		Full Demographic Adjusted		
Score	**Score**	**Percentile Rank**	**T-Score**	**Percentile Rank**	**Qualitative Descriptor**
Auditory Memory	82	12	42	21	Low Average
Visual Memory	80	9	40	16	Low Average
Immediate Memory	73	4	36	8	Mild Impairment
Delayed Memory	80	9	42	21	Low Average
Visual Working Memory	80	9	42	21	Low Average
Logical Memory I	7	16	43	24	Low Average
Logical Memory II	7	16	43	24	Low Average
Verbal Paired Associates I	8	25	47	38	Average
Verbal Paired Associates II	6	9	41	18	Low Average
Designs I	6	9	40	16	Low Average
Designs II	11	63	60	84	Above Average
Visual Reproduction I	3	1	32	4	Mild to Moderate Impairment
Visual Reproduction II	6	9	39	14	Mild Impairment
Spatial Addition	7	16	46	35	Average
Symbol Span	6	9	41	18	Low Average

Application of demographic-adjusted norms resulted in an increase in percentile rank across all measures; however, level of functioning did not change significantly. Most of the memory scores were in the low average range. Immediate memory was in the mild impairment range. WMS—IV subtests range from the mild to moderate impairment to above average range. All memory index scores were significantly lower than WAIS—IV GAI. The psychologist interpreted the memory testing as indicative of mild memory impairments compared to intellectual functioning, particularly on free recall for visual information.

Based on the demographically adjusted norms, the psychologist concluded that Ms. J. had average intellectual abilities and in some areas above average cognitive functioning. Based on her significantly lower

processing speed, working memory, and general memory functioning, the clinician inferred that she probably had some brain injury caused by the recent seizure event. This might affect her ability to work in a job requiring memory or quickly processing information but overall, her intellectual functioning would not preclude her from working. The psychologist also recommended that she might consider some vocational training or a 2-year college program.

The Case Study 2 example is proposed as a misuse of demographically adjusted norms. Why would this be a misapplication? There is a clear significant medical event which potentially could result in a change in brain functioning. The examinee wants to work and the demographic adjustments indicate that she is functioning in the average range on most intellectual skills. In fact, she has above average skills in some domains that perhaps warrant more educational investment.

The fundamental question when applying demographic adjustments to norms is "*do these variables accurately represent the patients pre-morbid intellectual ability?*" In this case, does female gender accurately represent the client's pre-morbid state? The answer would be yes. Is 12 years of education an accurate estimate of her pre-morbid ability? The answer is likely to be no because her ability to function in school and her ability to reach her potential was severely limited by her medical condition which also likely had a significant influence on her brain development. Based on her family history of high educational attainment, 12 years of education may have been the lower bounds of her potential. Does ethnicity equal African-American accurately reflect her pre-morbid abilities? The answer could be yes but it is probably no. As stated earlier in this chapter, a myriad of psychosocial factors likely related to the quality of education, access to appropriate healthcare, and socioeconomic status, may impact the test performance of African-Americans. Ms. J.'s background suggests she may not have experienced problems in these areas to the same degree as other African-Americans. The most appropriate scores to interpret in this case are probably age-adjusted normative scores. In that context, the clinician can determine how she is functioning on testing relative to people her age. These age adjusted scores reflect: (a) the developmental trajectory of her cognitive functioning in the context of a chronic seizure disorder, and (b) the possible adverse effects of her recent status epilepticus. The age-adjusted scores can be used to see if her cognition improves, remains stable, or worsens in the future.

Out of interest, if one wanted to see how her current scores relate to how she might have developed, in *the absence of a longstanding seizure disorder*, the best demographic adjustment to use for her might be 16 years of education with no adjustments for sex and ethnicity. Table 4.15 illustrates the changes in WAIS–IV performance when that model is applied.

TABLE 4.15 Ms. J.'s Age-Adjusted and Education Adjusted WAIS—IV Scores

	Age Adjusted		Education Adjusted		
Score	**Score**	**Percentile Rank**	**T-Score**	**Percentile Rank**	**Qualitative Descriptor**
Verbal Comprehension Index	89	23	33	5	Mild to Moderate Impairment
Perceptual Reasoning Index	88	21	39	14	Mild Impairment
Working Memory Index	86	18	35	7	Mild Impairment
Processing Speed Index	81	10	34	6	Mild to Moderate Impairment
Full Scale IQ	83	13	32	4	Mild to Moderate Impairment
General Ability Index	87	19	34	6	Mild to Moderate Impairment
Vocabulary	9	37	41	18	Low Average
Similarities	6	9	29	2	Moderate Impairment
Information	9	37	40	16	Low Average
Block Design	8	25	40	16	Low Average
Matrix Reasoning	7	16	37	10	Mild Impairment
Visual Puzzles	9	37	46	35	Average
Digit Span	7	16	35	7	Mild Impairment
Arithmetic	8	25	39	14	Mild Impairment
Coding	5	5	30	2	Mild to Moderate Impairment
Symbol Search	8	25	40	16	Low Average

Applying the education only adjustments changes the results dramatically. Where previously, scores were mostly average the scores are now mostly in the impaired to low average range. Does this suggest that she had a precipitous drop in cognitive functioning subsequent to the severe seizure event? Possibly, but more than likely, her cognitive functioning has been compromised over a long period of time due to the impact of the seizures and treatment during critical periods of brain development. There may be an additional impact of the status epilepticus but it may not be possible to determine any causality between the

event and her current level of functioning. Conceptually, if one considers her broader background, Ms. J. is an individual who has experienced a significant compromise to her intellectual functioning as a consequence of a chronic neurological condition. In chronic medical, psychiatric, and developmental disorders that directly affect the examinees ability to complete their education, education level is not a good proxy for pre-morbid ability. This is why in the previous case, the adjustments were appropriate. The patient's Parkinsonism did not affect his educational attainment therefore demographic adjustments were appropriate even though he had a chronic neurological condition. In the present case, age adjusted normative data are most appropriate.

The other issue with applying demographically adjusted norms in this case relates to applying the results to daily functioning. Demographic norms are not designed to infer how the person will function in the general community, in a job, or in an academic setting. For example, if an individual with 8 years of education scores a 90 on FSIQ, their education-adjusted score will be in the above average range. An individual with 18 years of education with an FSIQ of 105 will be in the low average range. Would you predict the individual with above average education-adjusted scores to be more successful in an academic environment that the individual with the low average score? Of course not, one examinee has already shown difficulty in school and the other success by their level of educational attainment. Education adjusted scores tell you how the individual compares to people who had similar difficulty or success with educational attainment. So to infer adequate cognitive capacity for job or academic success based on educational adjustments is an inappropriate use of those scores. The scores that should be applied are the age-adjusted scores.

SUMMARY

The significant relationship between cognitive test performance and background characteristics of the individual is well established. In particular, sex, education, and ethnicity have been identified as factors that need to be considered when interpreting performance on cognitive tests. Sex differences in cognitive functioning are typically very small with gender advantages observed in both directions (i.e., males perform better on some tests while females perform better on others). Education effects occur due to exposure to more information, better test-taking skills, and more training in basic skills such as reading and writing. Educational attainment is also affected by cognitive ability; therefore, individuals with better pre-morbid skills often attain more education. Racial/ethnic group differences exist due to socioeconomic and sociopolitical differences that

impact education quality, attainment, medical access, socioeconomic status, and other background factors that affect test performance. The differences are not due to how the tests are constructed or item/test bias. It is difficult to measure all the factors that produce racial/ethnic group differences, particularly the interaction of multiple complex forces, therefore race/ethnicity serves as a proxy for these effects.

The WAIS–IV shows large effects for education level, particularly when comparing the extremes of the education distribution. Sex effects on the WAIS–IV are small and ethnic group effects are moderate to large. The impact of demographic factors is significantly smaller for the WMS–IV. Applying demographic adjustments to normative data for the WAIS–IV and WMS–IV will have the greatest impact on WAIS–IV scores, particularly for patients at the ends of the demographic distribution (i.e., 8 years of education or 18 years of education). Therefore, the most clinical significance will be observed in patients at the extremes of the distribution.

The concept and clinical practice of adjusting normative data for demographics has been debated for decades. While there is no consensus about the application of such normative adjustments, many clinicians routinely use this methodology as part of neuropsychological practice. The purpose of using demographic adjustments is to *identify a change* in cognitive functioning from a pre-morbid level. Given that stated goal, it is very important to ascertain if the demographic adjustments are a good estimate of the patient's pre-morbid functioning. Demographic adjustments are not appropriate for use in individual's suffering chronic psychiatric, developmental, medical, or neurological conditions that have directly affected their educational attainment. Similarly, racial/ethnic group adjustments are not appropriate for all members of a specific group. The background factors that impact the group as a whole may or may not be experienced by the individual being assessed. Demographic adjustment to norms are not appropriate for making decisions about an individual's ability to function in general society, or in vocational or educational settings, and the norms do not necessarily reflect their capacity to understand court proceedings, financial functioning, or to consent for treatment. The clinician is solely responsible for application of normative adjustments and must have a rationale for applying the adjustments to a specific patient.

KEY LEARNING

- Patient background characteristics impact expected level of performance on cognitive tests.
- Sex, education, and race/ethnicity have well documented effects on cognitive test performance.

- Adjusting cognitive tests for background characteristics of the patient is controversial but has been a part of clinical neuropsychological practice for decades.
- Demographic adjustments are available for Education, and Education/Sex/Ethnicity for the WAIS–IV and WMS–IV as an ACS report option.
- WAIS–IV and WMS–IV demographic norms are derived using inferential norming techniques and are presented as T-scores.
- Demographic adjustments have a large impact on patients at the extremes of the distribution (education level = 8 or 18 years) but only small changes occur for individuals in the middle (education level = 12 or 13–15 years) of the distribution.
- Demographic adjustments are larger for WAIS–IV than WMS–IV and the level of adjustment varies on measures within WAIS–IV (e.g., VCI > PSI). It is not possible to use a rule of thumb or clinical estimation to accurately adjust scores.
- Demographic adjustments lower the correlation between variables resulting in more and larger discrepancies in performance (use significance level and base rate data).
- Demographic adjustments are appropriate for estimating if an examinees performance is unexpectedly low or high signifying a change in cognitive functioning.
- Demographic adjustments assume that the variables are a reasonable estimation of an examinees pre-morbid functioning.
- Demographic adjustments are not appropriate for individuals with chronic psychiatric, medical, developmental, or neurological conditions, if those conditions had a significant impact on the patient's educational attainment.
- Demographic adjustments should not be used for determining eligibility for intellectual disability or learning disability.
- Demographic adjustments should not be used to determine functional capacity in the general population, job, or educational setting unless specific circumstances warrant such use. It is incumbent upon the examiner to provide a rationale in these situations.
- Demographic norms should not be used to determine competency or death penalty eligibility (e.g., culpability/intellectual disability) unless the purpose is to establish whether brain injury, loss of cognitive functioning, or dementia is affecting the patient.
- It is the responsibility of the clinician to use demographic adjustments responsibly and with clear understanding of how adjustments affect scores and the implications of making those adjustments on the patient.
- Adjusted norms are presented as T-scores (mean = 50, SD = 10). The T-scores are interpreted in ACS using the descriptors listed below.

A clinician, however, can choose to use look-up tables for T-score conversions to percentile ranks and apply the same classification descriptors as are used for age-adjusted normative data (e.g., superior, high average, and extremely low):

- >55 "Above Average"
- 45–54 "Average"
- 40–44 "Low Average"
- 35–39 "Mild Impairment"
- 30–34 "Mild to Moderate Impairment"
- 25–29 "Moderate Impairment"
- 20–24 "Moderate to Severe Impairment"
- <20 "Severe Impairment".

References

Baker, D. P., Leon, J., Smith Greenaway, E., Collins, J., & Movit, M. (2011). The education effect on population health: a reassessment. *Population Development Review*, *37*, 307–332. doi:10.1111/j.1728-4457.2011.00412.x.

Belzunces dos Santos, E., de Souza Silva Tudesco, I., Caboclo, L. O. S. F., & Yacubian, E. M. T. (2011). Low educational level effects on the performance of healthy adults on a neuropsychological protocol suggested by the commission on neuropsychology of the Liga Brasileira de Epilepsia. *Arquivos Neuropsychiatrica*, *65*, 778–784.

Bernard, L. C. (1989). Halstead-Reitan neuropsychological test performance of black, hispanic, and white young adult males from poor academic backgrounds. *Archives of Clinical Neuropsychology*, *4*, 267–274.

Boone, K. B., Victor, T. L., Wen, J., Ranzani, J., & Ponton, M. (2007). The association between neuropsychological scores and ethnicity, language, and acculturation variables in a large patient population. *Archives of Clinical Neuropsychology*, *22*, 355–365.

Brickman, A. M., Cabo, R., & Manly, J. J. (2006). Ethical issues in cross-cultural neuropsychology. *Applied Neuropsychology*, *13*, 91–100.

Byrd, D. A., Miller, S. W., Reilly, J., Weber, S., Wall, T. L., & Heaton, R. K. (2006). Early environmental factors, ethnicity, and adult cognitive test performance. *The Clinical Neuropsychologist*, *20*, 243–260.

Camarata, S., & Woodcock, R. (2006). Sex differences in processing speed: developmental effects in males and females. *Intelligence*, *34*, 231–252.

Dori, G. A., & Chelune, G. J. (2004). Education-stratified base-rate information on discrepancy scores within and between the Wechsler adult intelligence scale–third edition and the Wechsler memory scale–third edition. *Psychological Assessment*, *16*, 146–154.

Golden, C. J., & van den Broek, A. (1998). Potential impact of age- and education-corrected scores on HRNB score patterns in participants with focal brain injury. *Archives of Clinical Neuropsychology*, *13*, 683–694.

Gonzalez, H. M., Whitfield, K. E., West, B. T., Williams, D. R., Lichtenberg, P. A., & Jackson, J. S. (2007). Modified-symbol digit modalities test for African Americans, caribbean black americans, and non-latino whites: nationally representative normative data from the National Survey of American Life. *Archives of Clinical Neuropsychology*, *22*, 605–613.

Gonzalez, R., Grant, I., Miller, S. W., Taylor, M. J., Schweinsburg, B. C., Carey, C. L., et al. (2006). Demographically adjusted normative standards for new indices of performance

on the paced auditory serial addition task (PASAT). *The Clinical Neuropsychologist*, *20*, 396–413.

Gorsuch, R. L. (2003, August). Update on continuous norming. Paper presented at the annual meeting of the american psychological association, Toronto, Canada. Heaton, R. K. & Grain, I. (1992). *Comprehensive norms for an expanded halstead-reitan battery: A supplement for the WAIS-R*. Odessa, FL: Psychological Assessment.

Gur, R. C., Richard, J., Calkins, M. E., Chiavacci, R., Hansen, J. A., Bilker, W. B., et al. (2012). Age group and sex differences in performance on a computerized neurocognitive battery in children age 8–21. *Neuropsychology*, *26*, 251–265.

Heaton, R. K., Avitable, N., Grant, I., & Matthews, C. G. (1999). Further cross validation of regression-based neuropsychological norms with an update for the Boston Naming Test. *Journal of Clinical and Experimental Neuropsychology*, *21*, 572–582.

Heaton, R. K., Grant, I., & Matthews, C. G. (1991). *Comprehensive norms for and expanded halstead-reitan battery: Demographic corrections, research findings, and clinical applications*. Odessa, FL: Psychological Assessment Resources.

Heaton, R. K., Taylor, M. J., & Manly, J. J. (2003). Demographic effects and the use of demographically corrected norms with the WAIS–III and WMS–III. In D. S. Tulsky, D. H. Saklofske, G. J. Chelune, R. K. Heaton, R. J. Ivnik, R. Bornstein, A. Prifitera, & M. F. Ledbetter (Eds.), Clinical interpretation of the WAIS–III and WMS–III. San Diego: Academic Press.

Herlitz, A., Airaksinen, E., & Nordstrom, E. (1999). Sex differences in episodic memory; the impact of verbal and visuospatial ability. *Neuropsychology*, *13*, 590–597.

Jefferson, A. L., Gibbons, L. E., Rentz, D. M., Carvalho, J. O., Manly, J., Bennett, D. A., et al. (2011). A life course model of cognitive activities, socioeconomic status, education, reading ability, and cognition. *Journal of the American Geriatric Society*, *59*, 1403–1411. doi:10.1111/j.1532-5415.2011.03499.x.

Karzmark, P., Heaton, R. K., Grant, I., & Matthews, C. G. (1984). Use of demographic variables to predict overall level of performance on the Halstead-Reitan Battery. *Journal of Consulting and Clinical Psychology*, *52*, 663–665.

Kramer, J. H., Delis, D. C., Kaplan, E., O'Donnell, L., & Prifitera, A. (1997). Developmental differences in verbal learning. *Neuropsychology*, *11*, 577–584.

Krueger, L. E., & Salthouse, T. A. (2010). Differences in acquisition, not retention, largely contribute to sex differences in multitrial word recall performance. *Personality and Individual Differences*, **49**, 768–772.

Lejbak, L., Crossley, M., & Vrbancic, M. (2011). A male advantage for spatial and object but not verbal working memory using the n-back task. *Brain and Cognition*, *76*, 191–196.

Lewin, C., Wolgers, G., & Herlitz, A. (2001). Sex differences favoring women in verbal but no visuospatial episodic memory. *Neuropsychology*, *15*, 165–173.

Liu, Y., Julkunen, V., Paajanen, T., Westman, E., Wahlund, L. O., Aitken, A., et al. (2012). Education increases reserve against Alzheimer's disease-evidence from structural MRI analysis. *Neuroradiology*, *54*, 929–938 [Epub ahead of print].

Longman, R. S., Saklofske, D. H., & Fung, T. S. (2007). WAIS–III percentile scores by education and sex for U.S. and Canadian populations. *Assessment*, *14*, 426–432.

Lucas, J. A., Ivnik, R. J., Willis, F. B., Ferman, T. J., Smith, G. E., Parfitt, F. C., et al. (2005). Mayo's older African-Americans normative studies: Normative data for commonly used clinical neuropsychological measures. *The Clinical Neuropsychologist*, *19*, 162–183.

Manly, J. J. (2005). Advantages and disadvantages of separate norms for African Americans. *The Clinical Neuropsychologist*, *19*, 270–275.

Manly, J. J., & Echemendia, R. J. (2007). Race-specific norms: using the model of hypertension to understand issues of race, culture, and education in neuropsychology. *Archives of Clinical Neuropsychology*, *22*, 319–325.

McBain, R., Norton, D., & Chen, Y. (2010). A female advantage in basic face recognition is absent in schizophrenia. *Psychiatry Research*, *177*, 12–17.

McDaniel, M. A. (2006). Estimating state IQ: measurement challenges and preliminary correlates. *Intelligence*, *34*, 607–619.

Moering, R. G., Schinka, J. A., Mortimer, J. A., & Graves, A. B. (2004). Normative data for elderly african americans for the stroop color and word test. *Archives of Clinical Neuropsychology*, *19*, 61–71.

Morgan, E. E., Woods, S. P., Scott, J. C., Childers, M., Beck, J. M., Ellis, R. J., et al. HIV Neurobehavioral Research Center (HNRC) Group (2008). Predictive validity of demographically adjusted normative standards for the HIV Dementia Scale. *Journal of Clinical and Experimental Neuropsychology*, *30*, 83–90.

Murray, A. D., Staff, R. T., McNeil, C. J., Salarirad, S., Ahearn, T. S., Mustafa, N., et al. (2011). The balance between cognitive reserve and brain imaging biomarkers of cerebrovascular and Alzheimer's diseases. *Brain*, *134*, 3687–3696. [Epub 2011 Nov 18].

Norman, M. A., Evans, J. D., Miller, S. W., & Heaton, R. K. (2000). Demographically corrected norms for the california verbal learning test. *Journal of Clinical and Experimental Neuropsychology*, *22*, 80–94.

Pearson (2009). Advanced clinical solutions for the WAIS–IV/WMS–IV*: Technical manual*. San Antonio, TX: NCS Pearson.

Ragland, J. D., Coleman, A. R., Gur, R. C., Glahn, D. C., & Gur, R. E. (2000). Sex differences in brain-behavior relationships between verbal episodic memory and resting regional cerebral blood flow. *Neuropsychologia*, **38**, 451–461.

Randolph, C., Lansing, A. E., Ivnik, R. J., Cullum, C. M., & Hermann, B. P. (1999). Determinants of confrontation naming performance. *Archives of Clinical Neuropsychology*, *14*, 489–496.

Reitan, R. M., & Wolfson, D. (1996). Differential relationships of age and education to WAIS subtest scores among brain-damaged and control groups. *Archives of Clinical Neuropsychology*, *11*, 303–311.

Roid, G. H. (2003). *Stanford-Binet intelligence scales, fifth edition, technical manual*. Itasca, IL: Riverside.

Sharp, E. S., & Gatz, M. (2011). Relationship between education and dementia: an updated systematic review. *Alzheimer Disease and Associated Disorders*, *24*, 289–304.

Taylor, M. J., & Heaton, R. K. (2001). Sensitivity and specificity of the WAIS–III/WMS–III demographically corrected factor scores in neuropsychological assessment. *Journal of the International Neuropsychological Society*, *7*, 867–874.

Tombaugh, T. N. (2004). Trail making test A and B: normative data stratified by age and education. *Archives of Clinical Neuropsychology*, *19*, 203–214.

Tombaugh, T. N., Kozak, J., & Rees, L. (1999). Normative data stratified by age and education for two measures of verbal fluency: FAS and animal naming. *Archives of Clinical Neuropsychology*, *14*, 167–177.

Trahan, D. E., & Quintana, J. W. (1990). Analysis of gender effects upon verbal and visual memory performance in adults. *Archives of Clinical Neuropsychology*, *5*, 325–334.

van der Sluis, S., Posthum, D., Dolan, C. V., de Geus, E. J. C., Colom, R., & Boomsma, D. I. (2006). Sex differences on the dutch WAIS-III. *Intelligence*, *34*, 273–289.

Wechsler, D. (2007). *Wechsler memory scale* (*3rd ed.*). San Antonio, TX: The Psychological Corporation.

Weirich, S., Hoffmann, F., Meissner, L., Heinz, A., & Bengner, T. (2011). Sex influence on face recognition memory moderated by presentation duration and reencoding. *Neuropsychology*, *25*, 806–813.

Weiss, L. G., Chen, H., Harris, J. G., Holdnack, J. A., & Saklofske, D. H. (2010). WAIS–IV use in societal context. In L. G. Weiss, D. H. Saklofske, D. L. Coalson, & S. E. Raiford

(Eds.), *WAIS–IV clinical use and interpretation: Scientist–practitioner perspectives*. San Diego: Academic Press.

Weiss, L. G., Harris, J. G., Prifitera, A., Courville, T., Rolfhus, E., Saklofske, D. H., et al. (2006). WISC–IV interpretation in societal context. In L. G. Weiss, D. H. Saklofske, A. Prifitera, & J. A. Holdnack (Eds.), WISC–IV advanced clinical interpretation. San Diego: Academic Press.

Wilkins, C., Rolfhus, E., Weiss, L. & Zhu, J. J. (2005). A new method for calibrating translated tests with small sample sizes. Paper presented at the 2005 annual meeting of the American Educational Research Association, Montreal, Canada.

Zahodne, L. B., Glymour, M. M., Sparks, C., Bontempo, D., Dixon, R. A., MacDonald, S. W., et al. (2011). Education does not slow cognitive decline with aging: 12-year evidence from the Victoria longitudinal study. *Journal of the International Neuropsychological Society*, *17*, 1039–1046 [Epub 2011 Sep 19].